New Distance Measure of Single-Valued Neutrosophic Sets and Its Application

Han-Liang Huang*

School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, People's Republic of China

A single-valued neutrosophic set (SVNS) is an instance of a neutrosophic set, which can be used to handle uncertainty, imprecise, indeterminate, and inconsistent information in real life. In this paper, a new distance measure between two SVNSs is defined by the full consideration of truth-membership function, indeterminacy-membership function, and falsity-membership function for the forward and backward differences. Then the similarity measure, the entropy measure, and the index of distance are also presented. Finally, two illustrative examples are given to demonstrate the effectiveness of the proposed clustering method and multicriteria decision-making method based on the distance (similarity) measure between SVNSs. © 2016 Wiley Periodicals, Inc.

1. INTRODUCTION

Uncertainty, imprecise, incomplete, and inconsistent information can be found in many real-life systems and may enter some problems in a much more complex ways. The theory of fuzzy set (FS) proposed by Zadeh[1] in 1965 has achieved a great success in various real applications to handle uncertainty. The traditional FS uses one real value $\mu_A(x) \in [0, 1]$ to represent the grade of membership of a FS A defined on a universe X. Sometimes, it may be very difficult to assign the membership value for a FS. And we need to consider the truth membership as well as the falsity membership for proper description of an object in some real-life probelms. The intuitionistic fuzzy set (IFS) on a universe X was introduced by Atanassov in 1983.[2] It is a generalization of FS, which has a degree of membership $\mu_A(x) \in [0, 1]$ and a degree of nonmembership $\nu_A(x) \in [0, 1]$ such that $\mu_A(x) + \nu_A(x) \leq 1$ for each $x \in X$. By Gau and Buehrer, the vague set (VS) was defined in 1993.[3] But it was proven that the VS is the same as IFS by Bustine and Burillo.[4] Both FS and IFS cannot handle the indeterminate and inconsistent information which exists in a belief system. Smarandache[5] originally introduced the concept of neutrosophic set (NS), which is a branch of philosophy and is a mathematical tool for studying the origin, nature, and scope of neutralities. A NS is represented independently by a truth-membership function $T_A(x)$, an indeterminacy-membership function $I_A(x)$, and a

*Author to whom all correspondence should be addressed; e-mail: hl_huang1980.student@ sina.com.

falsity-membership function $F_A(x)$, where $T_A(x)$, $I_A(x)$ and $F_A(x)$ are real standard or nonstandard subsets of $]^-0, 1^+[$ (see Definition 1). Although NS generalizes the concepts of FS, IFS, and others from philosophical point of view, it will be difficult to apply in real scientific and engineering areas. Wang et al.[6] introduced a single-valued neutrosophic set (SVNS) (see Definition 2), which can handle the problems involving imprecise, indeterminacy, and inconsistent data. Since SVNS is a powerful general formal framework, it has received considerable attention for many researchers. The information energy and correlation of SVNSs were presented by Ye in Ref. 7. Then some authors considered the decision-making method by the use of SVNSs.[7–11]

Clustering plays an important part in analyzing the real world, such as data mining, pattern recognition, machine learning, and so on. Over the past few decades, researchers used different kinds of tools to study the clustering methods. Ruspini[12] first proposed a fuzzy clustering approach by defining the concept of fuzzy division. Under the intuitionistic fuzzy environment, the clustering analysis method is comprehensively studied by using the method of the IFSs. Based on the traditional hierarchical clustering procedure, the intuitionistic fuzzy aggregation operator and the basic distance measures between IFSs, an intuitionistic fuzzy hierarchical algorithm is introduced for clustering IFSs by Xu.[13] Then Xu[14] developed an intuitionistic fuzzy C-means algorithm to cluster IFSs based on the well-known fuzzy C-means clustering method[15] and the basic distance measures between IFSs. Zhang et al.[16] and Xu et al.[17] also proposed the clustering algorithms for IFSs by defining the association coefficients and the similarity measures of IFSs. As the information aggregation operators involved in the above literature are mainly $\bigvee$ and $\bigwedge$, Lin and Huang[18] generalized the operators of IFSs to t-norm and t-conorm, respectively, and improved the clustering method presented in Ref. 16. Then Huang[19] extend the algorithms to cluster interval-valued intuitionistic fuzzy sets (IVIFSs) by introducing the notions of (T,S)-composition matrix and (T,S)- interval-valued intuitionistic fuzzy equivalence matrix. Recently, Ye[20] studied the single-valued neutrosophic clustering algorithm to deal with data represented by SVNSs. But the distance measure defined in Ref. 20 only considered the absolute difference between each truth-membership function, indeterminacy-membership function, and falsity-membership function of two SVNSs. From a comprehensive consideration of the problem, we extend the distance measure of SVNSs in this paper.

The paper is organized as follows: In Section 2, some basic concepts are reviewed. In Section 3, the notions of the distance measure, the similarity measure, the entropy measure, and the index of distance are introduced. In Section 4, three illustrate examples are given to show the effectiveness of the new distance measure applied in clustering and decision making. Finally, a conclusion is provided in Section 5.

2. PRELIMINARIES

The NS is a part of neutrosophy, which studies the origin, nature, and scope of neutralities. It is the generalization of the concepts of fuzzy sets, interval-valued fuzzy set, intuitionisitic fuzzy set.

DEFINITION 1.[5] *Let X be a space of points (objects), with a generic element in X denoted by x. A neutrosophic set A in X is characterized by a truth-membership function $T_A(x)$, an indeterminacy-membership function $I_A(x)$, and a falsity-membership function $F_A(x)$. The functions $T_A(x)$, $I_A(x)$, and $F_A(x)$ are real standard of non-standard subsets of $]^-0, 1^+[$. That is, $T_A x : x \rightarrow]^-0, 1^+[$, $I_A x : x \rightarrow]^-0, 1^+[$, and $T_A x : x \rightarrow]^-0, 1^+[$. Thus, there is no restriction on the sum of $T_A(x)$, $I_A(x)$, and $F_A(x)$, so $^-0 \leq \sup T_A(x) + \sup I_A(x) + \sup F_A(x) \leq 3^+$.*

Owing to the difficulty of applying the NS in real scientific and engineering, Wang et al. proposed the concepts of SVNS. It is a subclass of a NS.

DEFINITION 2.[6] *Let X be a space of points (objects) with generic elements in X denoted by x. A SVNS A in X is characterized by a truth-membership function $T_A(x)$, an indeterminacy-membership function $I_A(x)$, and a falsity-membership function $F_A(x)$. For each point x in X, $T_A(x), I_A(x), F_A(x) \in [0, 1]$. Therefore, a SVNS A can be denoted by $A = \{< x, T_A(x), I_A(x), F_A(x) > | x \in X\}$.*

The following expressions are defined by Wang et al. in Ref. 6 for SVNSs A, B:

(1) $A \subseteq B$ if and only if $T_A(x) \leq T_B(x), I_A(x) \geq I_B(x), F_A(x) \geq F_B(x)$ for any x in X.
(2) $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.
(3) $A^c = \{< x, F_A(x), 1 - I_A(x), T_A(x) > | x \in X\}$, where A^c denotes the complement of A.

Ye[20] defined the generalized single-valued neutrosophic weighted distance measure as

$$d_p(A, B) = \left\{ \frac{1}{3} \sum_{i=1}^{n} w_i [|T_A(x_i) - T_B(x_i)|^p \right.$$

$$\left. + |I_A(x_i) - I_B(x_i)|^p + |F_A(x_i) - F_B(x_i)|^p] \right\}^{1/p} \tag{1}$$

where A, B are two SVNSs, $p > 0, w_i(i = 1, 2, \ldots, n)$ are the weight of the element $x_i(i = 1, 2, \ldots, n)$ with $w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$.

Remark 1. The distance measure (1) only consider the absolute difference between each truth-membership function, indeterminacy-membership function, and falsity-membership function. But sometimes we need to think about the reaction of the falsity-membership function or indeterminacy-membership function. This may cause one-sidedness in some practical problems we considered. For example, A player completes his (her) action in a match. We use the truth-membership function to express the successful action and use the falsity-membership function to express the failure action. At the end of the match, the score of the player should be considered by the positive effects of the successful action and the negative impact of the failure action. In this condition, Equation 1 is not suitable to distinguish the differences between several players.

Example 1. Suppose there are three players whose performances are represented by three SVNSs $A = \langle 0.9, 0, 0.3 \rangle$, $B = \langle 0.7, 0, 0.1 \rangle$, and $C = \langle 0.7, 0, 0.5 \rangle$. Obviously, the distance between A and B is more closer than the distance between A and C. But by Equation 1, we have $d(A, B) = d(A, C)$, which is not consistent with our intuition.

3. DISTANCE MEASURE AND SIMILARITY MEASURE

This section proposes a new distance measure and similarity measure between two SVNSs.

DEFINITION 3. *For two SVNSs A and B in a universe of discourse* $X = \{x_1, x_2, \ldots, x_n\}$, *which are denoted by* $A = \{< x_i, T_A(x_i), I_A(x_i), F_A(x_i) > |x_i \in X\}$ *and* $B = \{< x_i, T_B(x_i), I_B(x_i), F_B(x_i) > |x_i \in X\}$. *The single-valued neutrosophic weighted distance measure are defined by*

$$d_\lambda(A, B) = \left[\sum_{j=1}^{n} w_j \left(\sum_{i=1}^{4} \beta_i \varphi_i(x_j) \right)^\lambda \right]^{1/\lambda} \tag{2}$$

where

$$\lambda > 0, \quad \beta_i \in [0, 1] \quad and \quad \sum_{i=1}^{4} \beta_i = 1, \quad w_j \in [0, 1] \quad and \quad \sum_{j=1}^{n} w_j = 1,$$

$$\varphi_1(x_j) = \frac{|T_A(x_j) - T_B(x_j)|}{3} + \frac{|I_A(x_j) - I_B(x_j)|}{3} + \frac{|F_A(x_j) - F_B(x_j)|}{3},$$

$$\varphi_2(x_j) = max \left\{ \frac{2 + T_A(x_j) - I_A(x_j) - F_A(x_j)}{3}, \frac{2 + T_B(x_j) - I_B(x_j) - F_B(x_j)}{3} \right\}$$

$$- min \left\{ \frac{2 + T_A(x_j) - I_A(x_j) - F_A(x_j)}{3}, \frac{2 + T_B(x_j) - I_B(x_j) - F_B(x_j)}{3} \right\},$$

$$\varphi_3(x_j) = \frac{|T_A(x_j) - T_B(x_j) + I_B(x_j) - I_A(x_j)|}{2},$$

$$\varphi_4(x_j) = \frac{|T_A(x_j) - T_B(x_j) + F_B(x_j) - F_A(x_j)|}{2}.$$

PROPOSITION 1. *The distance measure* $d_\lambda(A, B)$ *for* $\lambda > 0$ *satisfies the following properties:*

(D1) $0 \leq d_\lambda(A, B) \leq 1$;

(D2) $d_\lambda(A, B) = 0$ *if and only if* $A = B$;

(D3) $d_\lambda(A, B) = d_\lambda(B, A)$;

($D4$) If $A \subseteq B \subseteq C$, C is an SVNS in X, then $d_\lambda(A, B) \le d_\lambda(A, C)$, and
$d_\lambda(B, C) \le d_\lambda(A, C)$.

Proof.

- (D1) It is easy to see that $\varphi_i(x_j) \in [0, 1]$ $(i = 1, 2, 3, 4)$, $(j = 1, 2, \ldots, n)$, so $0 \le d_\lambda(A, B) \le 1$ hold.
- (D2) When $A = B$, we can have $\varphi_i(x_j) = 0$, $(i = 1, 2, 3, 4)$, $(j = 1, 2, \ldots, n)$. So $d_\lambda(A, B) = 0$. On the other hand, when $d_\lambda(A, B) = 0$ hold, we can see that $\varphi_i(x_j) = 0$, $(i = 1, 2, 3, 4)$, $(j = 1, 2, \ldots, n)$. It implies $T_A(x_j) = T_B(x_j)$, $I_A(x_j) = I_B(x_j)$ and $F_A(x_j) = F_B(x_j)$, that is $A = B$.
- (D3) Straightforward.
- (D4) When $A \subseteq B \subseteq C$, it has $T_A(x_j) \le T_B(x_j) \le T_C(x_j)$, $I_A(x_j) \ge I_B(x_j) \ge I_C(x_j)$ and $F_A(x_j) \ge F_B(x_j) \ge F_C(x_j)$ for $j = 1, 2, \ldots, n$. Then we have

$$|T_A(x_j) - T_B(x_j)| \le |T_A(x_j) - T_C(x_j)|, |T_B(x_j) - T_C(x_j)| \le |T_A(x_j) - T_C(x_j)|,$$

$$|I_A(x_j) - I_B(x_j)| \le |I_A(x_j) - I_C(x_j)|, |I_B(x_j) - I_C(x_j)| \le |I_A(x_j) - I_C(x_j)|,$$

$$|F_A(x_j) - F_B(x_j)| \le |F_A(x_j) - F_C(x_j)|, |F_B(x_j) - F_C(x_j)| \le |F_A(x_j) - F_C(x_j)|,$$

$$\frac{2 + T_A(x_j) - I_A(x_j) - F_A(x_j)}{3} \le \frac{2 + T_B(x_j) - I_B(x_j) - F_B(x_j)}{3}$$

$$\le \frac{2 + T_C(x_j) - I_C(x_j) - F_C(x_j)}{3},$$

$$0 \le \frac{T_B(x_j) - T_A(x_j) + I_A(x_j) - I_B(x_j)}{2} \le \frac{T_C(x_j) - T_A(x_j) + I_A(x_j) - I_C(x_j)}{2},$$

$$0 \le \frac{T_B(x_j) - T_A(x_j) + F_A(x_j) - F_B(x_j)}{2} \le \frac{T_C(x_j) - T_A(x_j) + F_A(x_j) - F_C(x_j)}{2},$$

Hence,

$$\varphi_i^{AB}(x_j) \le \varphi_i^{AC}(x_j), \;\; \varphi_i^{BC}(x_j) \le \varphi_i^{AC}(x_j), \;\; i = 1, 2, 3, 4, \;\; j = 1, 2, \ldots, n.$$

Therefore, $d_\lambda(A, B) \le d_\lambda(A, C)$ and $d_\lambda(B, C) \le d_\lambda(A, C)$.

$\square$

Remark 2. When we choose $\beta_1 = 1$, $\beta_2 = \beta_3 = \beta_4 = 0$, distance measure $d_\lambda(A, B)$ in (2) reduces to $d_p(A, B)$ in (1).

DEFINITION 4. *For two SVNSs A and B in a universe of discourse $X = \{x_1, x_2, \ldots, x_n\}$, which are denoted by $A = \{< x_i, T_A(x_i), I_A(x_i), F_A(x_i) > |x_i \in X\}$ and $B = \{< x_i, T_B(x_i), I_B(x_i), F_B(x_i) > |x_i \in X\}$. The single-valued neutrosophic weighted similarity measure are defined by*

$$\vartheta_\lambda(A, B) = 1 - d_\lambda(A, B). \tag{3}$$

PROPOSITION 2. *The similarity measure $\vartheta_\lambda(A, B)$ for $\lambda > 0$ satisfies the following properties:*

(S1) $0 \leq \vartheta_\lambda(A, B) \leq 1$;

(S2) $\vartheta_\lambda(A, B) = 1$ *if and only if* $A = B$;

(S3) $\vartheta_\lambda(A, B) = \vartheta_\lambda(B, A)$;

(S4) *If* $A \subseteq B \subseteq C$, C *is a SVNS in* X, *then* $\vartheta_\lambda(A, C) \leq \vartheta_\lambda(A, B)$, *and*

$\vartheta_\lambda(A, C) \leq \vartheta_\lambda(B, C)$.

Entropy is very important for measuring uncertain information. De Luca and Termini[21] first axiomatized non-probabilistic entropy. Then Szmidt and Kacprzyk[22] proposed a nonprobabilistic-type entropy measure for intuitionistic fuzzy sets. Since the results were formulated for fuzzy (intuitionistic fuzzy) sets, they are expressed for the SVNSs as follows:

DEFINITION 5. *Let E be a set-to-point mapping: E : SVNSs $\rightarrow$ [0, 1], Then E is an entropy measure if it satisfies the following conditions:*

(E1) *$E(A) = 0$ (minimum) if and only if A or A^c is a crisp set;*

(E2) *$E(A) = 1$ (maximum) if and only if $A = A^c$, i.e., $T_A(x_i) = F_A(x_i)$, $I_A(x_i) = 0.5$ for all $x_i \in X$;*

(E3) *$E(A) \leq E(B)$ if A is less fuzzy than B, i.e., if*

$T_A(x_i) \leq T_B(x_i)$, $F_B(x_i) \leq F_A(x_i)$ *for* $T_B(x_i) \leq F_B(x_i)$ *and* $I_A(x_i) = I_B(x_i) = 0.5$,

or

$T_A(x_i) \geq T_B(x_i)$, $F_B(x_i) \geq F_A(x_i)$ *for* $T_B(x_i) \geq F_B(x_i)$ *and* $I_A(x_i) = I_B(x_i) = 0.5$;

(E4) *$E(A) = E(A^c)$.*

PROPOSITION 3. *$\vartheta_\lambda(A, A^c)$ is an entropy measure for a SVNS A.*

Proof.

(E1) Assume that A is a crisp set, i.e., $T_A(x_i) = 1$, $I_A(x_i) = F_A(x_i) = 0$ or $F_A(x_i) = 1$, $T_A(x_i) = I_A(x_i) = 0$ for all $x_i \in X$. No matter in which cases (A or A^C is a crisp set), we have $\vartheta_\lambda(A, A^c) = 0$. On the other hand, let $\vartheta_\lambda(A, A^c) = 0$, we can easily get A or A^C is a crisp set.

(E2) By (S2) in Proposition 2, we can have $\vartheta_\lambda(A, A^c) = 1 \Leftrightarrow A = A^c$.

(E3) when A is less fuzzy than B, suppose that $T_A(x_i) \leq T_B(x_i)$, $F_B(x_i) \leq F_A(x_i)$ for $T_B(x_i) \leq F_B(x_i)$ and $I_A(x_i) = I_B(x_i) = 0.5$, then we can calculate that $\varphi_i^{A,A^c}(x_j) \geq \varphi_i^{B,B^c}(x_j)$ ($i = 1, 2, 3, 4$) for $x_j \in X$ by Equation (2). This implies $d_\lambda(A, A^c) \geq d_\lambda(B, B^c)$. Furthermore, by Equation 3, we have $\vartheta_\lambda(A, A^c) \leq \vartheta_\lambda(B, B^c)$. The other case can be also proved by the same way.

(E4) From (S3) in Proposition 2, we can prove (E4). $\square$

Remark 3. In some cases, we not only think about the distance between A and B, but also need to consider the distance between A and B^c. So we can define the index of distance for two SVNSs A and B as follows.

DEFINITION 6. *For two SVNSs A and B, the index of distance is defined by*

$$I_\lambda(A, B) = \frac{d_\lambda(A, B)}{d_\lambda(A, B^c)}. \tag{4}$$

PROPOSITION 4. *The index of distance $I_\lambda(A, B)$ for two SVNSs A and B satisfies the following properties:*

- (I1) $I_\lambda(A, B) = 0$ *if and only if $A = B$;*

- (I2) $I_\lambda(A, B) = 1$ *if and only if $d_\lambda(A, B) = d_\lambda(A, B^c)$;*

- (I3) $I_\lambda(A, B) \to +\infty, \quad i.e., \ A = B^c, \ this means A and B are completely different;*

- (I4) *when $A = B = B^c$, the entropy measure of A and B reaches its maximum value;*

- (I5) $I_\lambda(A, B) < 1$ *means compare with B^c, is more similar to B;*

- (I6) $I_\lambda(A, B) > 1$ *means compare with B^c, A is less similar to B.*

4.   PRACTICAL EXAMPLES

In this section, three examples are given to demonstrate the application of the proposed distance measure.

Example 2. In Example 1, it sounds more reasonable that the distance between players A and B should be more closer than the distance between players A and C. Now let $\beta_i = \frac{1}{4}$, $\lambda = 1$, using Equation 2 to calculate the distance between SVNSs A and B, A and C. Then we have $d(A, B) = 0.0583$, $d(A, C) = 0.1417$, which is consistent with our intuition.

4.1.   Clustering Method Based on the Distance (Similarity) Measure of SVNSs and an Example

In the process of clustering under a single-valued neutrosophic data environment, an expert usually provides his/her preferences with SVNSs. Then we can use the method given in Ref. 20 to classify the objects. The method is described as follows:

Step 1. By use of Equations 2 and 3, we can calculate the similarity measure degree of SVNSs. Then we have a similarity matrix $C = (s_{ij})_{m \times m}$, where $s_{ij} = s_{ji} = \vartheta_\lambda(A_i, A_j)$ for $i, j = 1, 2, \ldots, m$.

Step 2. The process of building the composition matrices is repeated until it holds that

$$C \rightarrow C^2 \rightarrow C^4 \rightarrow \cdots \rightarrow C^{2^k} = C^{2^{k+1}},$$

i.e., C^{2^k} is an equivalent matrix, where

$$C^2 = C \circ C = (s'_{ij})_{m \times m} = \max_k \{\min(s_{ik}, s_{kj})\}_{m \times m}$$

for $i, j = 1, 2, \ldots, m$.

Step 3. For the equivalent matrix $C^{2^k} \triangleq \bar{C} = (\bar{s}_{ij})_{m \times m}$, we can construct a α-cutting matrix $\bar{C}_\alpha = (\bar{s}_{ij}^\alpha)_{m \times m}$ of $\bar{C}$, where

$$\bar{s}_{ij}^\alpha = \begin{cases} 0, \bar{s}_{ij} < \alpha; \\ 1, \bar{s}_{ij} \geq \alpha, \end{cases} \tag{5}$$

for $j, j = 1, 2, \ldots, m$ and α is the confidence level with $\alpha \in [0, 1]$.

Step 4. Classify A_i by choosing deferent confidence level α. Line i and k of $\bar{C}_\alpha$ are called α-congruence if $\bar{s}_{ij}^\alpha = \bar{s}_{kj}^\alpha$ for all $j = 1, 2, \ldots, m$. Then A_i should fall into the same category as A_k.

In the following, a real example adapted from Ye[20] is employed to demonstrate the application and effectiveness of the proposed clustering method mentioned above.

Example 3. A car market is going to classify five different cars of A_m ($m = 1, 2, 3, 4, 5$). Every car has six evaluation factors (attributes): (i) x_1, fuel consumption; (ii) x_2, coefficient of friction; (iii) x_3, price; (iv) x_4, comfortable degree; (v) x_5, design; (vi) x_6, security coefficient. The characteristics of each car under the six attributes are represented by the form of SVNSs, and then the single-valued neutrosophic data are as follows:

	x_1	x_2	x_3	x_4	x_5	x_6
A_1	$\langle 0.3, 0.2, 0.5 \rangle$	$\langle 0.6, 0.3, 0.1 \rangle$	$\langle 0.4, 0.3, 0.3 \rangle$	$\langle 0.8, 0.1, 0.1 \rangle$	$\langle 0.1, 0.3, 0.6 \rangle$	$\langle 0.5, 0.2, 0.4 \rangle$
A_2	$\langle 0.6, 0.3, 0.3 \rangle$	$\langle 0.5, 0.4, 0.2 \rangle$	$\langle 0.6, 0.2, 0.1 \rangle$	$\langle 0.7, 0.2, 0.1 \rangle$	$\langle 0.3, 0.1, 0.6 \rangle$	$\langle 0.4, 0.3, 0.3 \rangle$
A_3	$\langle 0.4, 0.2, 0.4 \rangle$	$\langle 0.8, 0.2, 0.1 \rangle$	$\langle 0.5, 0.3, 0.1 \rangle$	$\langle 0.6, 0.1, 0.2 \rangle$	$\langle 0.4, 0.1, 0.5 \rangle$	$\langle 0.3, 0.2, 0.2 \rangle$
A_4	$\langle 0.2, 0.4, 0.4 \rangle$	$\langle 0.4, 0.5, 0.1 \rangle$	$\langle 0.9, 0.2, 0.0 \rangle$	$\langle 0.8, 0.2, 0.1 \rangle$	$\langle 0.2, 0.3, 0.5 \rangle$	$\langle 0.7, 0.3, 0.1 \rangle$
A_5	$\langle 0.5, 0.3, 0.2 \rangle$	$\langle 0.3, 0.2, 0.6 \rangle$	$\langle 0.6, 0.1, 0.3 \rangle$	$\langle 0.7, 0.1, 0.1 \rangle$	$\langle 0.6, 0.2, 0.2 \rangle$	$\langle 0.5, 0.2, 0.3 \rangle$

Let $\lambda = 2$, choosing the weight vectors $w_j = 1/6$ ($j = 1, 2, \ldots\ldots, 6$) and $\beta_i = 1/4$ ($i = 1, 2, 3, 4$), then we use the similarity measure (3) to classify the five different cars of A_m ($m = 1, 2, \ldots, 5$) by the single-valued neutrosophic clustering algorithms.

First, we utilize the distance measure formula (2) to calculate the distance measures between each pair of SVNSs A_m ($m = 1, 2, \ldots, 5$). The results are as

follows:

$$d(A_1, A_2) = 0.1271, \quad d(A_1, A_3) = 0.1217, \quad d_(A_1, A_4) = 0.1655,$$
$$d(A_1, A_5) = 0.2000, \quad d(A_2, A_3) = 0.1078, \quad d(A_2, A_4) = 0.1419,$$
$$d(A_2, A_5) = 0.1258, \quad d(A_3, A_4) = 0.1821, \quad d(A_3, A_5) = 0.1723,$$
$$d(A_4, A_5) = 0.1928.$$

So we construct the following similarity matrix:

$$C = \begin{bmatrix} 1 & 0.8729 & 0.8783 & 0.8345 & 0.8000 \\ 0.8729 & 1 & 0.8922 & 0.8581 & 0.8742 \\ 0.8783 & 0.8922 & 1 & 0.8179 & 0.8277 \\ 0.8345 & 0.8581 & 0.8179 & 1 & 0.8072 \\ 0.8000 & 0.8742 & 0.8277 & 0.8072 & 1 \end{bmatrix}.$$

Then by Step 2 in Section 4.1, we have

$$C^2 = \begin{bmatrix} 1 & 0.8783 & 0.8783 & 0.8581 & 0.8729 \\ 0.8783 & 1 & 0.8922 & 0.8581 & 0.8742 \\ 0.8783 & 0.8922 & 1 & 0.8581 & 0.8742 \\ 0.8581 & 0.8581 & 0.8581 & 1 & 0.8581 \\ 0.8729 & 0.8742 & 0.8742 & 0.8581 & 1 \end{bmatrix}.$$

Due to $C^2 \nsubseteq C$, i.e., C is not an equivalent matrix, we keep calculating.

$$C^4 = \begin{bmatrix} 1 & 0.8783 & 0.8783 & 0.8581 & 0.8742 \\ 0.8783 & 1 & 0.8922 & 0.8581 & 0.8742 \\ 0.8783 & 0.8922 & 1 & 0.8581 & 0.8742 \\ 0.8581 & 0.8581 & 0.8581 & 1 & 0.8581 \\ 0.8742 & 0.8742 & 0.8742 & 0.8581 & 1 \end{bmatrix}.$$

$$C^8 = \begin{bmatrix} 1 & 0.8783 & 0.8783 & 0.8581 & 0.8742 \\ 0.8783 & 1 & 0.8922 & 0.8581 & 0.8742 \\ 0.8783 & 0.8922 & 1 & 0.8581 & 0.8742 \\ 0.8581 & 0.8581 & 0.8581 & 1 & 0.8581 \\ 0.8742 & 0.8742 & 0.8742 & 0.8581 & 1 \end{bmatrix}.$$

It is clearly that $C^4 \nsubseteq C^2$, but $C^8 = C^4$. That is, C^4 is an equivalent matrix, denoted by $\bar{C}$.

Finally, choosing deferent confidence level α, we can construct a α-cutting matrix $\bar{C}_\alpha$ by Equation 5.

(I) Let $0 \leq \alpha \leq 0.8581$, $\bar{C}_\alpha = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$. By Step 4 in Section 4.1, A_m ($m = 1, 2, \ldots, 5$) can be divided into one category $\{A_1, A_2, A_3, A_4, A_5\}$.

(II) Let $0.8581 < \alpha \leq 0.8742$, $\bar{C}_\alpha = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$. Then the cars A_m ($m = 1, 2, \ldots, 5$) can be divided into two categories $\{A_1, A_2, A_3, A_5\}$, $\{A_4\}$.

(III) Let $0.8742 < \alpha \leq 0.8783$, $\bar{C}_\alpha = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Then the cars A_m ($m = 1, 2, \ldots, 5$) can be divided into three categories $\{A_1, A_2, A_3\}$, $\{A_4\}$, $\{A_5\}$.

(IV) Let $0.8783 < \alpha \leq 0.8922$, $\bar{C}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Then the cars A_m ($m = 1, 2, \ldots, 5$) can be divided into four categories $\{A_1\}$, $\{A_2, A_3\}$, $\{A_4\}$, $\{A_5\}$.

(V) Let $0.8922 < \alpha \leq 1$, $\bar{C}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Then the cars A_m ($m = 1, 2, \ldots, 5$) can be divided into five categories $\{A_1\}$, $\{A_2\}$, $\{A_3\}$, $\{A_4\}$, $\{A_5\}$.

Remark 4. In Ref. 16, only three situations are obtained by the clustering algorithm based on the similar measure of IFSs; however, we can obtain five situations by our proposed method. Therefore, our method has better accuracy in some clustering problems.

Remark 5. The clustering results using our similarity measures are also different from the results presented in Ref. 20. In Ref. 20, most classifications are the same as our results except dividing the cars into two categories $\{A_1, A_2, A_3, A_4\}$, $\{A_5\}$. The reason for this difference is that the distance measure (1) only considers the absolute difference between each truth-membership function, indeterminacy-membership function, and falsity-membership function and ignores the reaction of the falsity-membership function or indeterminacy-membership function.

4.2.   Multicriteria Decision Making

Example 4. A manufacturing company which wants to select the best global supplier according to the core competencies of suppliers.[10] Now suppose that there are a set of four suppliers $A = \{A_1, A_2, A_3, A_4\}$ whose core competencies are evaluated by means of the following four criteria $\{C_1, C_2, C_3, C_4\}$. C_1: the level of technology innovation, C_2: the control ability of flow, C_3: the ability of management, C_4: the level

of service. Then, the weight vector for the four criteria is $w = (0.3, 0.25, 0.25, 0.2)$. When the four possible alternatives with respect to the above four criteria are evaluated by the similar method from the expert, we can obtain the following single-valued neutrosophic decision matrix A:

$$A = \begin{bmatrix} \langle 0.5, 0.1, 0.3 \rangle & \langle 0.5, 0.1, 0.4 \rangle & \langle 0.7, 0.1, 0.2 \rangle & \langle 0.3, 0.2, 0.1 \rangle \\ \langle 0.4, 0.2, 0.3 \rangle & \langle 0.3, 0.2, 0.4 \rangle & \langle 0.9, 0.0, 0.1 \rangle & \langle 0.5, 0.3, 0.2 \rangle \\ \langle 0.4, 0.3, 0.1 \rangle & \langle 0.5, 0.1, 0.3 \rangle & \langle 0.5, 0.0, 0.4 \rangle & \langle 0.6, 0.2, 0.2 \rangle \\ \langle 0.6, 0.1, 0.2 \rangle & \langle 0.2, 0.2, 0.5 \rangle & \langle 0.4, 0.3, 0.2 \rangle & \langle 0.7, 0.2, 0.1 \rangle \end{bmatrix}$$

By applying Equation 2 the distance between an alternative $A_i (i = 1, 2, 3, 4)$ and the ideal alternative $A^* = (< 1, 0, 0 >, < 1, 0, 0 >, < 1, 0, 0 >, < 1, 0, 0 >)$ are as follows:

$$d(A_1, A^*) = 0.3211, \ \ d(A_2, A^*) = 0.3575,$$

$$d(A_3, A^*) = 0.3305, \ \ d(A_4, A^*) = 0.3754,$$

with $\lambda = 2$ and $\beta_i = 1/4$ ($i = 1, 2, 3, 4$). This implies that the ranking order of the four suppliers is A_1, A_3, A_2, and A_4. Therefore, the best supplier is A_1. This result is the same as using Equation 1.

5. CONCLUSION

SVNSs can be applied in addressing problems with uncertain, imprecise, incomplete and inconsistent information existing in real scientific and engineering applications. Based on the related research achievements in SVNSs, we defined a new distance measure. It is a generalization of the existing distance measures defined in Ref. 20. Then, we also defined a new similarity measure, an entropy measure, and an index of distance under the single-valued neutrosophic environment. Two illustrative examples demonstrated the application of the proposed clustering analysis method and decision-making method. Compared to the existing methods for IFSs in Ref. 16 and for SVNSs in Ref. 20, the final results in this paper show that our method is more precise and reliable.

References

1. Zadeh LA. Fuzzy sets. Inform Control 1965;8:338–353.
2. Atanassov K. Intuitionistic fuzzy sets. Fuzzy Sets Syst 1986;20:87–96.
3. Gau WL, Buehrer DJ. Vague sets. IEEE Trans Syst Man Cybernet 1993;23:610–614.
4. Bustince H, Burillo P. Vague sets are intiutionistic fuzzy sets. Fuzzy Sets Syst 1996;79:403–405.

5. Smarandache F. A unifying field in logics. neutrosophy: neutrosophic probability, set and logic. Rehoboth: American Research Press; 1999.

6. Wang H, Smarandache F, Zhang YQ, Sunderraman R. Single valued neutrosophic sets. Multispace Multistruct 2010:4:410–413.

7. Ye J. Multicriteria decision-making method using the correlation coefficient under single-value neutrosophic environment. Int J Gen Syst 2013;42:386–394.

8. Ridvan S. Multi-criteria neutrosophic decision making method based on score and accuracy funtions under neutrosophic environment. Eprint Arxiv, 2014.

9. Ye J. A multicriteria decision-making method using aggregation operators for simplified neutrosophic sets. J Intell Fuzzy Syst 2014;26:2459–2466.

10. Ye J. Single valued neutrosophic cross-entropy for multicriteria decision making problems. Appl Math Modell 2014;38:1170–1175.

11. Zhang H, Wang J, Chen X. Interval neutrosophic sets and their application in multicriteria decision making problems. International Journal of Intelligent Systems 2014;31(5):1–15.

12. Ruspini EH. A new approach to clustering. Inform Control 1969;15:22–32.

13. Xu ZS. Intuitionistic fuzzy hierarchical clustering algorithms. J Syst Eng Electron 2009;20:90–97.

14. Xu ZS, Wu JJ. Intuitionistic fuzzy c-means clustering algorithms. J Syst Eng Electron 2010;21:580–590.

15. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. New York: Plenum Press; 1981.

16. Zhang HM, Xu ZS, Chen Q. On clustering approach to intuitionistic fuzzy sets. Control Decis 2007;22:882–888.

17. Xu ZS, Chen J, Wu J. Clustering algorithm for intuitionistic fuzzy sets. Inf Sci 2008;178:3775–3790.

18. Lin ML, Huang HL. (T,S)-based intuitionistic fuzzy composite matrix and its application. Int J Appl Math Stat 2011;23 54–63.

19. Huang HL. (T,S)-based interval-valued intuitionistic fuzzy composition matrix and its application for clustering. Iran J Fuzzy Syst 2012;9:7–19.

20. Ye J. Clustering methods using distance-based similarity measures of single-valued neutrosophic sets. J Intell Syst 2014;23:379–389.

21. De Luca AS, Termini S. A definition of nonprobabilistic entropy in the setting of fuzzy sets theory. Inform Control 1972;20:301–312.

22. Szmidt E, Kacprzyk J. Entropy for intuitionistic fuzzy sets. Fuzzy Sets Syst 2001;118:467–477.